

PATENT COOPERATION TREATY

PCT

NOTIFICATION OF ELECTION

Date of mailing (day/month/year) 10 January 2001 (10.01.01)	ETATS-UNIS D'AMERIQUE in its capacity as elected Office
International application No. PCT/US00/04118	Applicant's or agent's file reference 32313-PCT
International filing date (day/month/year) 18 February 2000 (18.02.00)	Priority date (day/month/year) 19 February 1999 (19.02.99)
Applicant	
MCKEOWN, Kathleen, R. et al	

1. The designated Office is hereby notified of its election made:

in the demand filed with the International Preliminary Examining Authority on:

05 September 2000 (05.09.00)

in a notice effecting later election filed with the International Bureau on:

2. The election was

X

was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

<p>The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland</p> <p>Facsimile No.: (41-22) 740.14.35</p>	<p>Authorized officer</p> <p>R. E. Stoffel</p> <p>Telephone No.: (41-22) 338.83.38</p>
---	--

PATENT COOPERATION TREATY

09/913745 B2313
PCT

From the INTERNATIONAL BUREAU

PCT

**NOTIFICATION CONCERNING
SUBMISSION OR TRANSMITTAL
OF PRIORITY DOCUMENT**

(PCT Administrative Instructions, Section 411)

Date of mailing (day/month/year)
08 June 2000 (08.06.00)

To:

TANG, Henry
Baker Botts, LLP
30 Rockefeller Plaza
New York, NY 10112-0228
ETATS-UNIS D'AMERIQUE BAKER BOTT S L.L.P.

00 JUN 26 PM 12:37

Applicant's or agent's file reference
32313-PCT

IMPORTANT NOTIFICATION

International application No.
PCT/US00/04118

International filing date (day/month/year)
18 February 2000 (18.02.00)

International publication date (day/month/year)
Not yet published

Priority date (day/month/year)
19 February 1999 (19.02.99)

Applicant

THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK et al

- The applicant is hereby notified of the date of receipt (except where the letters "NR" appear in the right-hand column) by the International Bureau of the priority document(s) relating to the earlier application(s) indicated below. Unless otherwise indicated by an asterisk appearing next to a date of receipt, or by the letters "NR", in the right-hand column, the priority document concerned was submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b).
- This updates and replaces any previously issued notification concerning submission or transmittal of priority documents.
- An asterisk(*) appearing next to a date of receipt, in the right-hand column, denotes a priority document submitted or transmitted to the International Bureau but not in compliance with Rule 17.1(a) or (b). In such a case, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.
- The letters "NR" appearing in the right-hand column denote a priority document which was not received by the International Bureau or which the applicant did not request the receiving Office to prepare and transmit to the International Bureau, as provided by Rule 17.1(a) or (b), respectively. In such a case, the attention of the applicant is directed to Rule 17.1(c) which provides that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity, upon entry into the national phase, to furnish the priority document within a time limit which is reasonable under the circumstances.

<u>Priority date</u>	<u>Priority application No.</u>	<u>Country or regional Office or PCT receiving Office</u>	<u>Date of receipt of priority document</u>
19 Febr 1999 (19.02.99)	60/120,659	US	22 May 2000 (22.05.00)

<p>The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland</p> <p>Faxsimile No. (41-22) 740.14.35</p>	<p>Authorized officer Marc Salzman</p> <p>Telephone No. (41-22) 338.83.38</p>
---	--



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : G06F 17/10, 17/27, 15/00		A2	(11) International Publication Number: WO 00/49517 (43) International Publication Date: 24 August 2000 (24.08.00)
<p>(21) International Application Number: PCT/US00/04118</p> <p>(22) International Filing Date: 18 February 2000 (18.02.00)</p> <p>(30) Priority Data: 60/120,659 19 February 1999 (19.02.99) US</p> <p>(71) Applicant (<i>for all designated States except US</i>): THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; 116th Street and Broadway, New York, NY 10027 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (<i>for US only</i>): MCKEOWN, Kathleen, R. [US/US]; 20 Prospect Road, Wayne, NJ 07470 (US). BARZILAY, Regina [US/US]; 548 Riverside Drive, Apt. 4B, New York, NY 10027 (US).</p> <p>(74) Agents: TANG, Henry et al.; Baker Botts, LLP, 30 Rockefeller Plaza, New York, NY 10112-0228 (US).</p>		<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>Without international search report and to be republished upon receipt of that report.</i></p>	
<p>(54) Title: MULTI-DOCUMENT SUMMARIZATION SYSTEM AND METHOD</p> <p>(57) Abstract</p> <p>A summary for a collection of related documents can be generated by extracting phrases from the documents which include common focus elements. Phrase intersection analysis is then performed on the extracted phrases to generate a phrase intersection table, where identical or equivalent phrases are identified. Temporal processing on the phrases in the phrase intersection table is performed to remove ambiguous time references and to sort the phrases in a temporal sequence. Sentence generation is then used to combine the phrases in the phrase intersection table into a coherent summary.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

MULTI-DOCUMENT SUMMARIZATION SYSTEM AND METHOD**SPECIFICATION****Statement of Government Rights**

The United States Government may have certain rights to the
5 invention set forth herein pursuant to a grant by the National Science Foundation,
Contract No. IRI-96-18797.

Statement of Related Applications

This application claims the benefit of United States provisional patent
application, Serial No. 60/120,659, entitled "Information Fusion in the Context of
10 Multi-Document Summarization," which was filed on February 19, 1999.

Field of the Invention

The present invention relates generally to information summarization
and more particularly relates to systems and methods for generating a summary for a
set of multiple, related documents.

Background of the Invention

The amount of information available today drastically exceeds that of
any time in history. With the continuing expansion of the Internet, this trend will
likely continue well into the future. Often, people conducting research of a topic are
faced with information overload as the number of potentially relevant documents
20 exceeds the researchers ability to individually review each document. To address this
problem, information summaries are often relied on by researchers to quickly evaluate
a document to determine if it is truly relevant to the problem at hand.

Given the vast collection of documents available, there is interest in
developing and improving the systems and methods used to summarize information
25 content. For individual documents, domain-dependent template based systems and

domain-independent sentence extraction methods are known. Such known systems can provide a reasonable summary of a single document. However, these systems are not able to compare and contrast related documents in a document set to provide a summary of the collection.

5 The ability to summarize collections of documents containing related information is desirable to further expedite the research process. For example, for a researcher interested in news stories regarding a certain event, a summary of all documents from a given source, or multiple sources, would provide a valuable overview of the documents within the set. From such a summary, the researcher may
10 be able to extract the information desired, or at the very least, make an informed decision regarding the relevance of the set of documents. Therefore, there remains a need for systems and methods which can generate a summary of related documents in a document set.

Summary of the Invention

15 It is an object of the present invention to provide a system and method for generating a summary of a set of multiple, related documents.

It is a further object of the present invention to provide a system and method for generating a summary of a set of multiple, related documents which use paraphrasing rules to detect similarities in non-identical phrases in the documents.

20 A present method for generating a summary of related documents in a collection includes extracting phrases from the documents which have common focus elements. Phrase intersection analysis is performed on the extracted phrases to generate a phrase intersection table. Temporal processing can be performed on the phrases in the phrase intersection table to remove ambiguous temporal references and
25 to sort the phrases in a temporal sequence. Sentence generation is performed using the phrases in the phrase intersection table to generate the multidocument summary.

30 Preferably, the phrase intersection analysis operation can include representing the phrases in tree structures having root nodes and children nodes; selecting those tree structures with verb root nodes; comparing the selected root nodes to the other root nodes to identify identical nodes; applying paraphrasing rules to non-

identical root nodes to determine if non identical nodes are equivalent; and evaluating the children nodes of those tree structures where the parent nodes are identical or equivalent. The tree structure can take the form of a DSYNT tree structure. The paraphrasing rules can include one or more rules which are selected from the group 5 consisting of ordering of sentence components, main clause versus a relative clause, different syntactic categories, change in grammatical features, omission of an empty head, transformation of one part of speech to another, and semantically related words.

In an embodiment of the present method, the temporal processing includes time stamping phrases based on a first occurrence of the phrase in the 10 collection; substituting date certain references for ambiguous temporal references; ordering the phrases based on the time stamp; and inserting a temporal marker if a temporal gap between phrases exceeds a threshold value.

Preferably, a phrase divergence processing operation can also be performed to include phrases that signal changes in focus of the documents in the 15 collection.

Sentence generation can include mapping the phrases, represented in the tree structure, to an input format of a language generation engine and then operating the language generation engine.

A present system for generating a summary of a collection of related 20 documents includes a storage device for storing the documents in the collection, a lexical database, such as WordNet, and a processing subsystem operatively coupled to the storage device and the lexical database. The processing subsystem is programmed to perform multiple document summarization including: accessing the documents in the storage device; using the lexical database to extract phrases from the documents 25 with similar focus elements; performing phrase intersection analysis on the extracted phrases to generate a phrase intersection table; performing temporal processing on the phrases in the phrase intersection table; and performing sentence generation using the phrases in the phrase intersection table.

The methods described above can be encoded in the form of a 30 computer program stored in computer readable media, such as CD-ROM, magnetic storage and the like.

Brief Description of the Drawing

Further objects, features and advantages of the invention will become apparent from the following detailed description taken in conjunction with the accompanying figures showing illustrative embodiments of the invention, in which

5 Figure 1 is a flow chart illustrating the operation of the present multiple document summarization system;

Figure 2 is a flow chart of a phrase intersection processing operation in accordance with the system operation of Figure 1;

10 Figure 3 is a pictorial diagram of a DSYNT tree structure for an exemplary sentence;

Figure 4 is a flow chart of a temporal processing operation in accordance with the system operation of Figure 1;

Figure 5 is a simplified block diagram of an embodiment of the present multiple document summarization system.

15 Throughout the figures, the same reference numerals and characters, unless otherwise stated, are used to denote like features, elements, components or portions of the illustrated embodiments. Moreover, while the subject invention will now be described in detail with reference to the figures, it is done so in connection with the illustrative embodiments. It is intended that changes and modifications can
20 be made to the described embodiments without departing from the true scope and spirit of the subject invention as defined by the appended claims.

Detailed Description of Preferred Embodiments

Figure 1 is a flow chart which provides an overview of the operation of the present multiple document summarization system. Initially, a set of documents, in computer readable format and grouped by a common theme or domain, is presented to the summarization system. From the collection of documents, entities are identified and sentences are extracted from the documents which are relevant to the focus of the articles. Entities can be identified and extracted in a number of ways, such as by use of an information extraction engine. A suitable information extraction engine is
30 TALENT, which is available from International Business Machines, Inc. In step 100,

phrases are extracted from the documents which include terms that are present in at least two of the documents. In addition, divergent phrases, which may be indicative of contrasts in the documents, are also extracted from the document in step 110.

Following extraction, phrase intersection processing (step 120) and phrase divergence 5 processing (step 130) are performed to evaluate and compare the extracted phrases and determine whether such phrases should be included in the resulting multiple document summary. Since phrases are extracted from multiple documents and can include temporal references which are ambiguous when taken out of context from the original document, temporal processing (step 140) is performed on the phrases 10 selected for the summary. Finally, sentence generation (step 150) is used to transform selected phrases into a coherent summary.

Figure 2 is a flow chart which further illustrates steps that can be performed in connection with phrase intersection processing of step 120. The selected phrases are grammatically parsed and represented in a tree structure, such as a 15 DSYNT tree diagram, which is generally known in the art. An example of such a diagram is illustrated in Figure 3. The parse trees can be generated by a conventional grammatical parser, such as Collin's parser. The DSYNT tree structure is a way of representing the constituent dependencies resulting from a predicate-argument sentence structure. In the tree structure, each non-auxiliary word in the sentence has a 20 node which is connected to its direct dependents. Grammatical features of each word are also stored in the node. To facilitate subsequent comparisons, words in the nodes are kept in their canonical form.

Returning to Figure 2, those trees which have root nodes which are verbs are selected and used as the basis for comparison. Each such verb based tree is 25 compared against the other trees derived from the sentences extracted from the documents in the collection (step 220). A comparison is made to determine if two nodes are identical (step 230). If two nodes are identical, those nodes are added to an output tree (step 235) and the nodes are evaluated to determine if there are further nodes descending from the root node (step 240). Such further nodes are referred to as 30 children nodes. If children nodes are present (step 245), the comparison in step 230 is repeated for each of child node. If the analysis of the children nodes is complete at

step 240, a determination is made as to whether the trees with common root nodes represent a phrase intersection (step 250). For example, if there is commonality in the root node and at least two children nodes of that root node, that phrase can be deemed complete and added to a phrase intersection table (step 255). If no phrase intersection is detected at step 250, the next parent node is selected for processing (step 260) and control returns to step 230.

Returning to step 230, if two nodes are not identical, it is still possible for the nodes to be equivalent. To make this determination, the present method employs a set of paraphrasing rules to evaluate the nodes (step 265). Paraphrasing, which can be broadly defined as alternative ways a human speaker can choose to "say the same thing" by using linguistic knowledge, generally occurs at a "surface" level, e.g., it is achieved by using semantically related words and syntactic transformations.

In the case of a set of related documents, theme sentences of the documents will generally be close semantically. This limits the scope of different paraphrasing types to be evaluated. From an analysis of paraphrasing patterns evaluated through themes of a training corpus derived from TDT, the following non-exhaustive set of paraphrasing categories have been found to occur with the greatest frequency:

1. ordering of sentence components: "*Tuesday they met...*" and "*They met ... Tuesday*";
2. main clause vs. a relative clause: "...*a building was devastated by the bomb*" and "...*a building, devastated by the bomb*";
3. realization in different syntactic categories, e.g., classifier vs. apposition: "*Palestinian leader Arafat*" and "*Arafat, Palestinian leader*", "*Pentagon speaker*" and "*speaker from the Pentagon*";
4. change in grammatical features: active/passive, time, number. "...*a building was devastated by the bomb*" and "...*the bomb devastated a building*";
5. omission of an empty head: "*group of students*" and "*students*";
6. transformation from one part of speech to another: "*building devastation*" and "...*building was devastated*"; and

7. using semantically related words such as synonyms: "*return*" and "*alight*", "*regime*" and "*government*".

The categories presented are used as paraphrasing rules by the present methods. The majority of these categories, such as ordering, can be fully implemented
5 in an automatic way,. However, some of the rules can be only approximated to a certain degree in an automated system. For example, identification of similarity based on semantic relations between words depends on the scope of coverage of the thesaurus employed. Word similarity can be established using relationships such as synonymy, hyponymy/hypernymy, and meronymy/holonymy which are detectable
10 using the WordNet language database which is described in the article "WordNet: A lexical Database for English", by G.A. Miller, Communications of the ACM, Vol. 38, No. 11. pp. 39-41, November 1995.

If any of the included paraphrasing rules are satisfied for non-identical nodes, the nodes are deemed equivalent (step 270). Equivalent nodes are added to the
15 output tree (step 235) and processed in the same manner as identical nodes. If no paraphrasing rule is applicable to non-identical nodes, there is no phrase intersection with the current tree (step 280).

In addition to phrase intersection processing, which compares phrases for similarity, it is also desirable to perform phrase divergence processing (step 130),
20 which compares selected phrases for differences. Phrase divergence may indicate a critical change in the course of events through a set of related documents and would be worthy of inclusion in a summary. For example, a collection of articles regarding a plane crash could begin with a focus on the passengers as "survivors" and later refer to "casualties," "victims," "bodies" and the like, which signify a turning point in the
25 events described by the documents. WordNet can also be used in phrase divergence processing by evaluating focus relationships such antonymy (e.g., "happiness is opposite to sadness").

Once phrases are selected from the documents for the summary,
temporal processing can be performed to sequence the phrases and eliminate
30 ambiguous temporal references. The flow chart of Figure 4 illustrates an overview of the temporal processing operations performed in the present methods. Using a rule

that an event is assumed to have occurred on the day that it is first reported, a time stamp can be applied to the selected phrases based on the earliest occurrence of the phrase in the collection of documents (step 405). In certain cases, phrases may include ambiguous temporal references, such as today, yesterday, etc. In this case,
5 such ambiguous references can be replaced by a date certain reference, such as by changing "Yesterday it was reported...." to "On 01/02/2000, it was reported...". Such substitutions, which are performed in step 410, can be implemented using the Emacs "calendar" package.

The extracted phrases can then be ordered in accordance with the
10 assigned date stamp (step 415). In certain cases, a large temporal gap may exist between consecutive phrases. In such a case, if the gap exceeds a threshold, such as two days, a temporal marker can be inserted between the phrases to indicate this gap in time (step 420). This may be significant, for example, in the case of a collection of news articles where the gap in time can also correspond to a change in focus in the
15 articles.

With the phrases selected and sorted in temporal order, sentence generation (step 150) can be performed to synthesize a coherent summary. Sentence generation involves two major operations. First, the DSYNT representation of the phrases to be used in sentence generation are mapped to the appropriate syntax of a
20 selected language generation engine. Then, the language generation engine is operated to arrange the phrases into coherent sentences. A suitable language generation engine is FUF/SURGE, which is available from Columbia University, New York, New York, as well as from Ben Gurion University, Department of Computer Science, Beer-Sheva, Israel. The acronym FUF stands for Functional Unification
25 Formalism interpreter and the acronym SURGE stands for syntactic realization grammar for text generation. The input specification for the FUF/SURGE engine includes a semantic role, *circumstantial*, which itself includes a temporal feature. The inclusion of the semantic attributes enables FUF/SURGE to perform various paraphrasing operations to the input phrases to improve the resulting sentences.

30 Figure 5 is a simplified block diagram of a multiple document summarization system in accordance with the present invention. The system 500

includes a processor section 505 wherein the processing operations set forth in Figure 1 are performed. The system also includes non-volatile storage coupled to the processor section 505 for document storage 510, collection summary storage 515, lexical database storage 518 and program storage 520. Generally these storage
5 systems are read/write data storage systems, such as magnetic media and read/write optical storage media. However, the document collection storage may take the form of read-only storage, such as a CD-ROM storage device. The system further includes RAM memory 525 coupled to the processor section for temporary storage during operation. The system 500 will generally include one or more input device 530 such
10 as a keyboard, digitizer, mouse and the like, which is coupled to the processor section 505. Similarly, a conventional display device 535 is generally provided which is also operatively coupled to the processor section.

The particular hardware embodiment is not critical to the practice of the present invention. Various computer platforms and architectures can be used to
15 implement the system 500, such as personal computers, workstations, networked computers, and the like. The functions described in the system can be performed locally or in a distributed manner, such as over a local area network or the Internet. For example, the document collection storage 510 may be at a remote archive location which is accessed by the processor section 505 via a connection to the Internet.

20 Although the present invention has been described in connection with specific exemplary embodiments, it should be understood that various changes, substitutions and alterations can be made to the disclosed embodiments without departing from the spirit and scope of the invention as set forth in the appended claims.

CLAIMS

1. A method for generating a summary of a plurality of related documents in a collection comprising:

5 extracting phrases having focus elements from the plurality of documents;

performing phrase intersection analysis on the extracted phrases to generate a phrase intersection table;

10 performing temporal processing on the phrases in the phrase intersection table; and

15 performing sentence generation using the phrases in the phrase intersection table.

2. The method of generating a summary as defined by claim 1, wherein the phrase intersection analysis comprises:

representing the phrases in tree structures having root nodes and children nodes;

15 selecting those tree structures with verb root nodes;

comparing the selected root nodes to the other root nodes to identify identical nodes;

20 applying paraphrasing rules to non-identical root nodes to determine if non identical nodes are equivalent; and

evaluating the children nodes of those tree structures where the parent nodes are identical or equivalent.

3. The method of claim 2, wherein the tree structure is a DSYNT tree structure.

25 4. The method of claim 2, wherein the paraphrasing rules are selected from the group consisting of ordering of sentence components, main clause versus a relative clause, different syntactic categories, change in grammatical features, omission of an

empty head, transformation of one part of speech to another, and semantically related words.

5. The method of claim 1, wherein the temporal processing includes:

time stamping phrases based on a first occurrence of the phrase in the
5 collection;

substituting date certain references for ambiguous temporal references;
ordering the phrases based on the time stamp; and
inserting a temporal marker if a temporal gap between phrases exceeds
a threshold value.

10 6. The method of claim 1, further comprising a phrase divergence processing
operation.

7. The method of claim 1, wherein the sentence generation includes mapping
phrases to an input format of a language generation engine and operating the language
generation engine.

15 8. A system for generating a summary of a plurality of related documents in a
collection comprising:

a storage device for storing the documents in the collection;
a lexical database; and
a processing subsystem, the processing subsystem being operatively
20 coupled to the storage device and the lexical database, the processing subsystem being
programmed to access the documents in the storage device and:

using the lexical database to extract phrases having focus elements
from the plurality of documents;

25 performing phrase intersection analysis on the extracted phrases to
generate a phrase intersection table;

performing temporal processing on the phrases in the phrase
intersection table; and

performing sentence generation using the phrases in the phrase intersection table.

9. The system for generating a summary as defined by claim 9, wherein the phrase intersection analysis processing further comprises:

5 representing the phrases as data structures having root nodes and children nodes;

selecting those data structures with verb root nodes;

comparing the selected root nodes to the other root nodes to identify identical nodes;

10 applying paraphrasing rules to non-identical root nodes to determine if non identical nodes are equivalent; and

evaluating the children nodes of those tree structures where the parent nodes are identical or equivalent.

15 10. The system of claim 9, wherein the data structure is a DSYNT tree structure.

11. The system of claim 9, wherein the paraphrasing rules are selected from the group consisting of ordering of sentence components, main clause versus a relative clause, different syntactic categories, change in grammatical features, omission of an empty head, transformation of one part of speech to another, and semantically related words.

20 12. The system of claim 8, wherein the temporal processing includes:

time stamping phrases based on a first occurrence of the phrase in the collection;

substituting date certain references for ambiguous temporal references;

25 ordering the phrases based on the time stamp; and

inserting a temporal marker if a temporal gap between phrases exceeds a threshold value.

13. The system of claim 8, further comprising a phrase divergence processing operation.
14. The system of claim 8, wherein the processing subsystem includes a language generation engine and wherein sentence generation includes mapping phrases to an input format of the language generation engine and then operating the language generation engine.
5
15. The system of claim 8, wherein the storage device for storing the documents in the collection is remotely located from the processing subsystem.
16. A computer readable media for programming a computer system to perform a method of generating a summary of a plurality of related documents in a collection comprising:
 - 10 extracting phrases having focus elements from the plurality of documents;
 - 15 performing phrase intersection analysis on the extracted phrases to generate a phrase intersection table;
 - 20 performing temporal processing on the phrases in the phrase intersection table; and
 - 25 performing sentence generation using the phrases in the phrase intersection table.
17. The computer readable media of claim 16, wherein the phrase intersection analysis comprises:
 - representing the phrases in tree structures having root nodes and children nodes;
 - selecting those tree structures with verb root nodes;
 - 25 comparing the selected root nodes to the other root nodes to identify identical nodes;
 - applying paraphrasing rules to non-identical root nodes to determine if non identical nodes are equivalent; and

evaluating the children nodes of those tree structures where the parent nodes are identical or equivalent.

18. The computer readable media of claim 17, wherein the tree structure is a
5 DSYNT tree structure.

19. The computer readable media of claim 17, wherein the paraphrasing rules are selected from the group consisting of ordering of sentence components, main clause versus a relative clause, different syntactic categories, change in grammatical features, omission of an empty head, transformation of one part of speech to another, and
10 semantically related words.

20. The computer readable media of claim 16, wherein the temporal processing includes:

time stamping phrases based on a first occurrence of the phrase in the collection;

15 substituting date certain references for ambiguous temporal references; ordering the phrases based on the time stamp; and inserting a temporal marker if a temporal gap between phrases exceeds a threshold value.

21. The computer readable media of claim 16, further comprising a phrase divergence processing operation.
20

22. The computer readable media of claim 16, wherein the sentence generation includes mapping phrases to an input format of a language generation engine and operating the language generation engine.

1/5

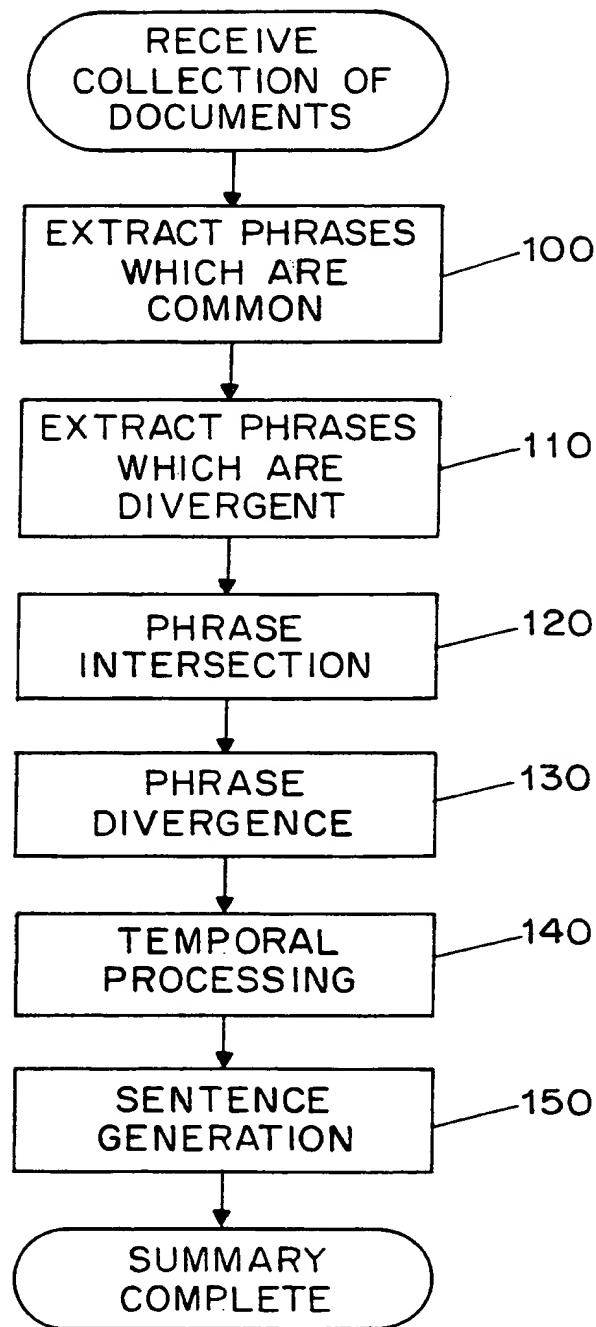
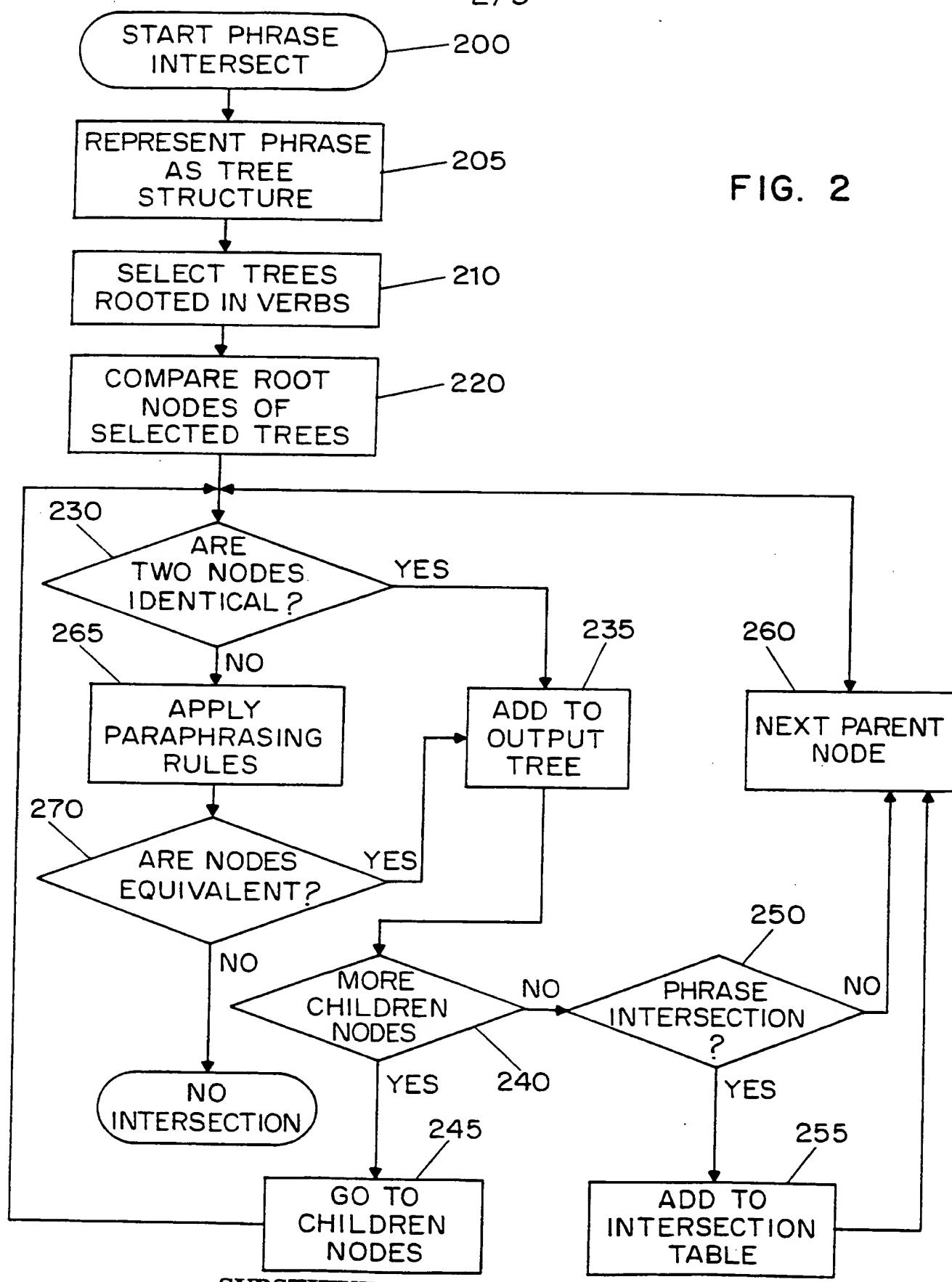


FIG. 1

2/5



SUBSTITUTE SHEET (RULE 26)

FIG. 2

09/913745

3/5

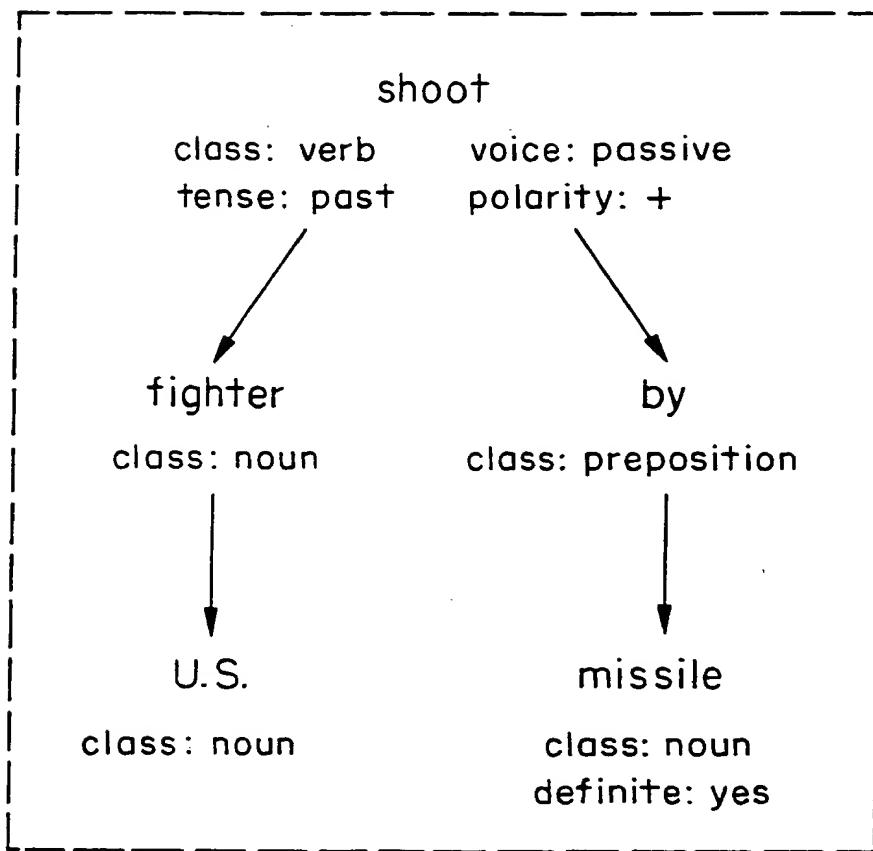


FIG. 3

09/913745

4/5

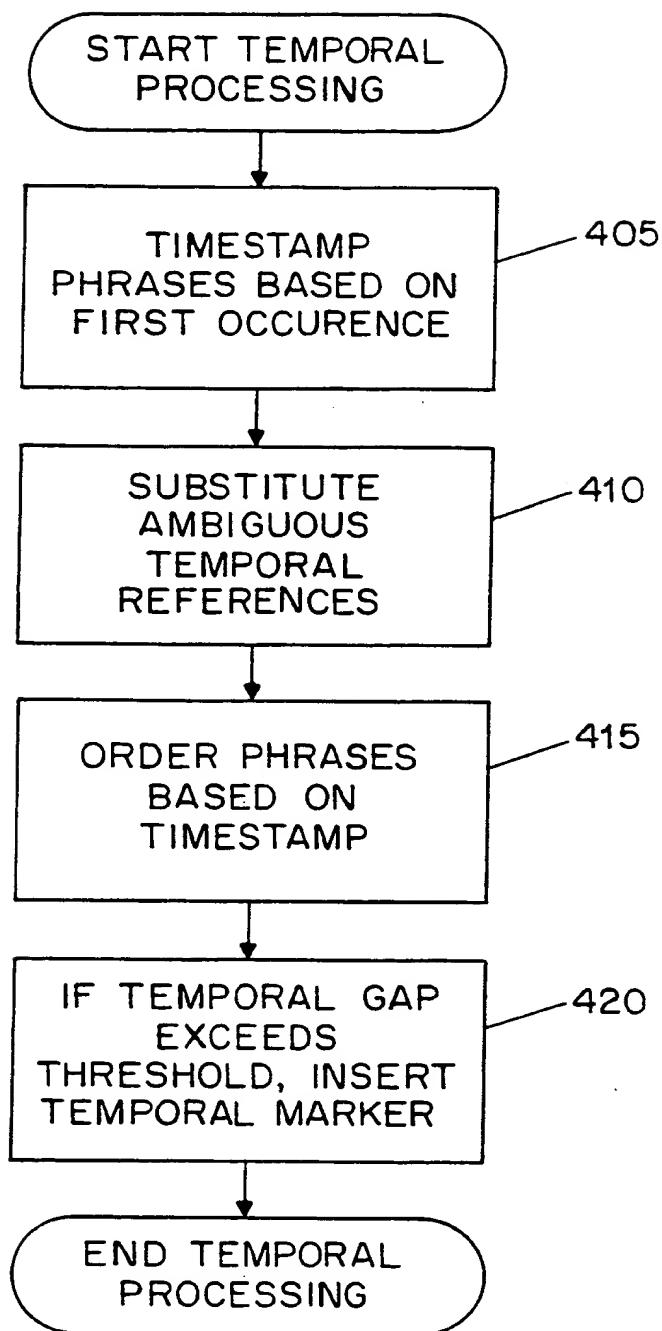


FIG. 4

09/913745

5/5

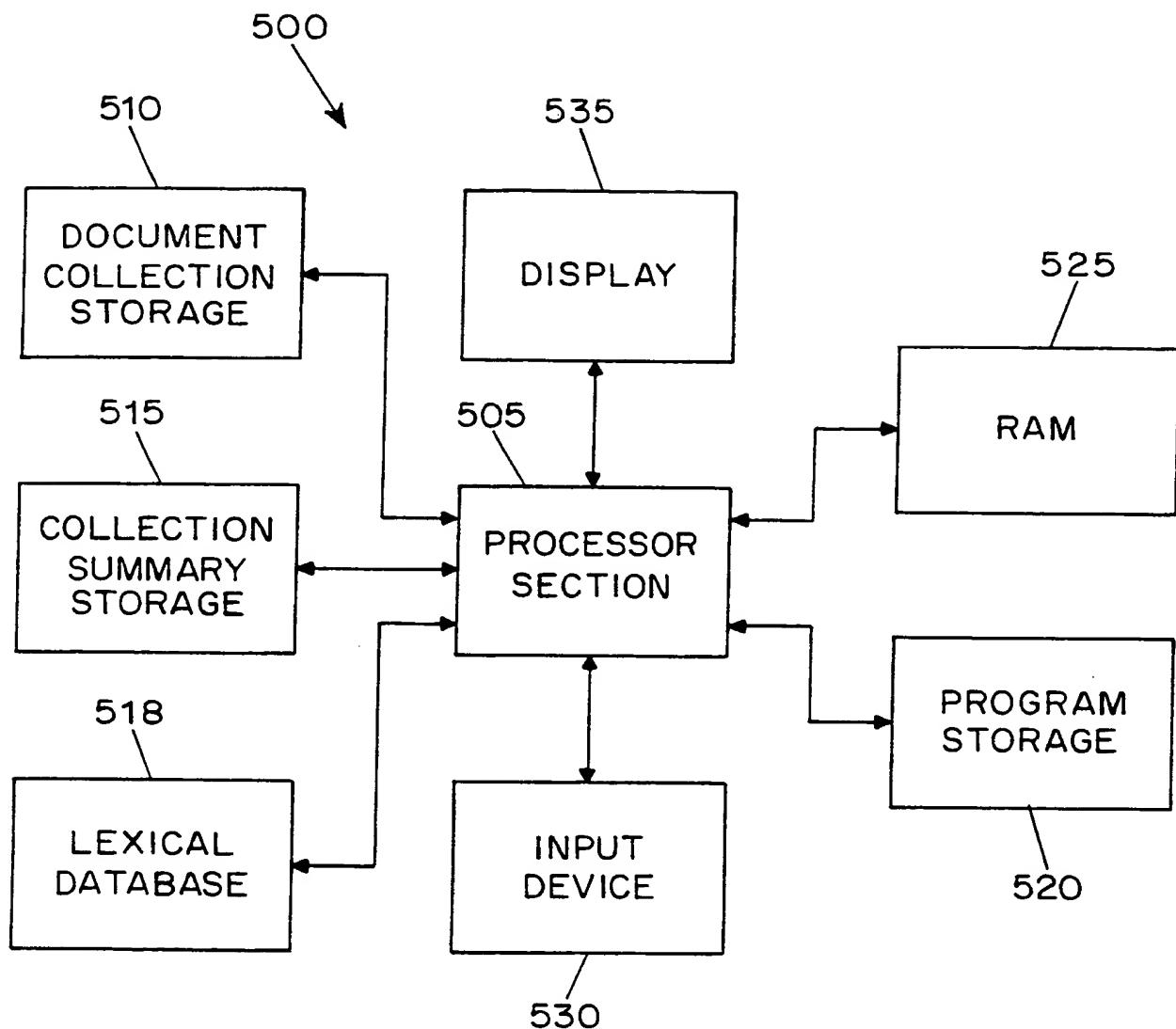


FIG. 5

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
24 August 2000 (24.08.2000)

PCT

(10) International Publication Number
WO 00/49517 A3

(51) International Patent Classification⁷: G06F 17/10, 17/27, 15/00

(21) International Application Number: PCT/US00/04118

(22) International Filing Date: 18 February 2000 (18.02.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 60/120,659 19 February 1999 (19.02.1999) US

(71) Applicant (for all designated States except US): THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; 116th Street and Broadway, New York, NY 10027 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): MCKEOWN, Kathleen, R. [US/US]; 20 Prospect Road, Wayne, NJ 07470 (US). BARZILAY, Regina [US/US]; 548 Riverside Drive, Apt. 4B, New York, NY 10027 (US).

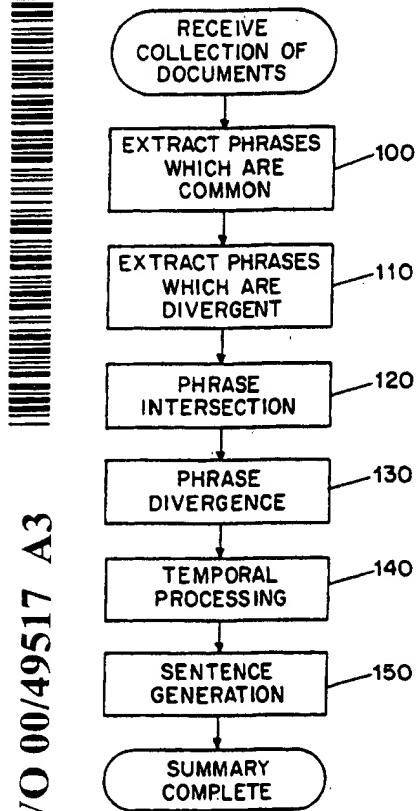
(74) Agents: TANG, Henry et al.; Baker Botts, LLP, 30 Rockefeller Plaza, New York, NY 10112-0228 (US).

(81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent

[Continued on next page]

(54) Title: MULTI-DOCUMENT SUMMARIZATION SYSTEM AND METHOD



(57) Abstract: A summary for a collection of related documents can be generated by extracting phrases (100) from the documents which include common focus elements. Phrase intersection analysis (120) is then performed on the extracted phrases (100) to generate a phrase intersection table, where identical or equivalent phrases are identified. Temporal processing (140) on the phrases in the phrase intersection table is performed to remove ambiguous time references and to sort the phrases in a temporal sequence. Sentence generation is then used to combine the phrases in the phrase intersection table into a coherent summary.

WO 00/49517 A3



(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent
(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,
MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM,
GA, GN, GW, ML, MR, NE, SN, TD, TG).

(88) Date of publication of the international search report
30 November 200

Published:

— With international search report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/04118

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/10, 17/27, 15/00

US CL : 704/1, 9, 10; 707/530, 531, 532

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 704/1, 9, 10; 707/530, 531, 532

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
A,P	US 5,924,108 A (FEIN ET AL) 13 JULY 1999, ABSTRACT.	1-22
A	US 5,638,543 A (PEDERSEN ET AL) 10 JUNE 1997.	1-22
Y	US 5,077,668 A (DOI) 31 DECEMBER 1991, ABSTRACT.	1, 8 AND 16
A	US 5,848,191 A (CHEN ET AL) 08 DECEMBER 1998	1-22
A	US 4,965,763 A (ZAMORA) 23 OCTOBER 1990.	1-22
Y	US 5,838,323 A (ROSE ET AL) 17 NOVEMBER 1998, ABSTRACT.	1, 8 AND 16
A	US 5,297,027 A (MORIMOTO ET AL) 22 MARCH 1994.	1-22

 Further documents are listed in the continuation of Box C. See patent family annex.

Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"		document defining the general state of the art which is not considered to be of particular relevance
"B"	"X"	earlier document published on or after the international filing date
"L"		document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reasons (as specified)
"O"	"Y"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"P"	"A"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
		document member of the same patent family

Date of the actual completion of the international search

03 JULY 2000

Date of mailing of the international search report

11 AUG 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

Patrick N. Edouard

Telephone No. (703) 308-6725

Joni Hill

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/04118

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,689,716 A (CHEN ET AL) 18 NOVEMBER 1997, ABSTRACT.	1, 8 AND 16
A	US 5,778,397 A (KUPIEC ET AL) 07 JULY 1998.	1-22
A	US 5,384,703 A (WITHGOTT ET AL) 24 JANUARY 1995.	1-22

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04118

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST/EAST

search term: (summarS or abstractS or condens) same (document or text or information)

09/913745

JC03 Rec'd PCT/PTO 16 AUG 2001
32313-PCT 070050.1161

IN THE INTERNATIONAL BUREAU OF WIPO

PCT RECEIVING OFFICE

Applicants	:	The Trustees of Columbia University in the City of New York et al.
International Application No.	:	PCT/US00/04118
International Filing Date	:	18 February 2000
Title of Invention	:	MULTI-DOCUMENT SUMMARIZATION SYSTEM AND METHOD

VIA AIRMAIL

COMMUNICATION
REQUESTING CHANGE OF APPLICANT'S NATIONALITY

International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20
Switzerland

Sir or Madam:

Under Rule 92^{bis}, Applicants' attorneys respectfully request that the nationality of the co-applicant Barzilay, Regina be corrected to IL (Israel). Ms. Barzilay is a citizen of Israel. Enclosed is a substitute Sheet No. 2 of the PCT Request as well as a substitute Sheet No. 1 of the Demand reflecting the correction.

Applicants' attorneys request that this correction be made of record, and that the International Bureau notify the elected states.

09/913745

JC03 Rec'd PCT/TTO 16 AUG 2001

32313-PCT 070050.1161

Respectfully submitted,

BAKER BOTTS, LLP

Dated: October 31, 2000


Henry Tang
Reg. No. 29,705

Paul D. Ackerman
Reg. No. 39,891

Attorneys for Applicant
(212) 705-5000

Encls.

The demand must be filed directly with competent International Preliminary Examining Authority or, if two or more Authorities are with the one chosen by the applicant. The full name or two-letter code of that Authority may be indicated by the applicant on the line IPEA/ US

PCT

CHAPTER II

DEMAND

under Article 31 of the Patent Cooperation Treaty:

The undersigned requests that the international application specified below be the subject of international preliminary examination according to the Patent Cooperation Treaty and hereby elects all eligible States (except where otherwise indicated).

For International Preliminary Examining Authority use only

Identification of IPEA		Date of receipt of DEMAND
Box No. I IDENTIFICATION OF THE INTERNATIONAL APPLICATION		
International application No. PCT/US00/04118	International filing date (day/month/year) 18 February 2000 (18.02.00)	Applicant's or agent's file reference 32312-PCT (Earliest) Priority date (day/month/year) 19 February 1999 (19.02.99)
Title of invention MULTI-DOCUMENT SUMMARIZATION SYSTEM AND METHOD		
Box No. II APPLICANT(S)		
Name and address: (<i>Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.</i>) THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK 116th Street and Broadway New York, NY 10027 US		Telephone No.: Facsimile No.: Teleprinter No.:
State (<i>that is, country</i>) of nationality: US	State (<i>that is, country</i>) of residence: US	
Name and address: (<i>Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.</i>) MCKEOWN, KATHLEEN R. 20 Prospect Road Wayne, NJ 07470 US		
State (<i>that is, country</i>) of nationality: US	State (<i>that is, country</i>) of residence: US	
Name and address: (<i>Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.</i>) BARZILAY, REGINA 548 Riverside Drive, Apt. 4B New York, NY 10027 US		
State (<i>that is, country</i>) of nationality: IL	State (<i>that is, country</i>) of residence: US	
<input type="checkbox"/> Further applicants are indicated on a continuation sheet.		

Continuation of Box No. III FURTHER APPLICANTS AND/OR (FURTHER) INVENTOR(S)*If none of the following sub-boxes is used, this sheet is not to be included in the request.*

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

BARZILAY, REGINA
548 Riverside Drive, Apt. 4B
New York, NY 10027
US

This person is:

applicant only
 applicant and inventor
 inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality:
IL

State (that is, country) of residence:
US

This person is applicant all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box for the purposes of:

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

This person is:

applicant only
 applicant and inventor
 inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality:

State (that is, country) of residence:

This person is applicant all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box for the purposes of:

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

This person is:

applicant only
 applicant and inventor
 inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality:

State (that is, country) of residence:

This person is applicant all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box for the purposes of:

Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)

This person is:

applicant only
 applicant and inventor
 inventor only (If this check-box is marked, do not fill in below.)

State (that is, country) of nationality:

State (that is, country) of residence:

This person is applicant all designated States all designated States except the United States of America the United States of America only the States indicated in the Supplemental Box for the purposes of:

Further applicants and/or (further) inventors are indicated on another continuation sheet.

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

09/913745
BAKER BOTTS L.L.P.

PCT 00 DEC 27 PM 2:36

To: HENRY TANG BAKER BOTTS, LLP 30 ROCKEFELLER PLAZA NEW YORK, NY 10112 0228
--

**NOTIFICATION OF RECEIPT
OF DEMAND BY COMPETENT INTERNATIONAL
PRELIMINARY EXAMINING AUTHORITY**

(PCT Rules 59.3(e) and 61.1(b), first sentence
and Administrative Instructions, Section 601(a))

21 DEC 2000

Applicant's or agent's file reference 32313-PCT	IMPORTANT NOTIFICATION	
International application No. PCT/US00/04118	International filing date (day/month/year) 18 FEB 00	Priority date (day/month/year) 19 FEB 99
Applicant THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF & NEW YORK		

1. The applicant is hereby notified that this International Preliminary Examining Authority considers the following date as the date of receipt of the demand for international preliminary examination of the international application:

05 September 2000 (05-09-00)

2. That date of receipt is:

- the actual date of receipt of the demand by this Authority (Rule 61.1(b)).
- the actual date of receipt of the demand on behalf of this Authority (Rule 59.3(e)).
- the date on which this Authority has, in response to the invitation to correct defects in the demand (Form PCT/IPEA/404), received the required corrections.

3. **ATTENTION:** That date of receipt is AFTER the expiration of 19 months from the priority date. Consequently, the election(s) made in the demand does (do) not have the effect of postponing the entry into the national phase until 30 months from the priority date (or later in some Offices) (Article 39(1)). Therefore, the acts for entry into the national phase must be performed within 20 months from the priority date (or later in some Offices) (Article 22). For details, see the *PCT Applicant's Guide*, Volume II.

(If applicable) This notification confirms the information given by telephone, facsimile transmission or in person on:

ON 8/15/00

8/15/00

4. Only where paragraph 3 applies, a copy of this notification has been sent to the International Bureau.

Name and mailing address of the IPEA/ Assistant Commissioner for Patent Box PCT Washington, D.C. 20231 Attn: RO/US Facsimile No. 703-305-3230	Authorized officer <i>M. Jay</i> Telephone No. <i>703. 308-0454</i> <i>303. 6445</i>
---	---

09/913745

32313

RGT

From the INTERNATIONAL BUREAU

PCT

**NOTICE INFORMING THE APPLICANT OF THE
COMMUNICATION OF THE INTERNATIONAL
APPLICATION TO THE DESIGNATED OFFICES**

(PCT Rule 47.1(c), first sentence)

Date of mailing (day/month/year) 24 August 2000 (24.08.00)		To: TANG, Henry Baker Botts, LLP 30 Rockefeller Plaza New York, NY 10112-0228 ETATS-UNIS D'AMERIQUE	
Applicant's or agent's file reference 32313-PCT		IMPORTANT NOTICE	
International application No. PCT/US00/04118	International filing date (day/month/year) 18 February 2000 (18.02.00)	Priority date (day/month/year) 19 February 1999 (19.02.99)	
Applicant THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK et al			

1. Notice is hereby given that the International Bureau has communicated, as provided in Article 20, the international application to the following designated Offices on the date indicated above as the date of mailing of this Notice:
AU,KP,KR,US

In accordance with Rule 47.1(c), third sentence, those Offices will accept the present Notice as conclusive evidence that the communication of the international application has duly taken place on the date of mailing indicated above and no copy of the international application is required to be furnished by the applicant to the designated Office(s).

2. The following designated Offices have waived the requirement for such a communication at this time:

AE,AL,AM,AP,AT,AZ,BA,BB,BG,BR,BY,CA,CH,CN,CR,CU,CZ,DE,DK,DM,EA,EE,EP,ES,FI,GB,GD,
GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,
NO,NZ,OA,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZW
The communication will be made to those Offices only upon their request. Furthermore, those Offices do not require the applicant to furnish a copy of the international application (Rule 49.1(a-bis)).

3. Enclosed with this Notice is a copy of the international application as published by the International Bureau on 24 August 2000 (24.08.00) under No. WO 00/49517

REMINDER REGARDING CHAPTER II (Article 31(2)(a) and Rule 54.2)

If the applicant wishes to postpone entry into the national phase until 30 months (or later in some Offices) from the priority date, a demand for international preliminary examination must be filed with the competent International Preliminary Examining Authority before the expiration of 19 months from the priority date.

It is the applicant's sole responsibility to monitor the 19-month time limit.

Note that only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination.

REMINDER REGARDING ENTRY INTO THE NATIONAL PHASE (Article 22 or 39(1))

If the applicant wishes to proceed with the international application in the national phase, he must, within 20 months or 30 months, or later in some Offices, perform the acts referred to therein before each designated or elected Office.

For further important information on the time limits and acts to be performed for entering the national phase, see the Annex to Form PCT/IB/301 (Notification of Receipt of Record Copy) and Volume II of the PCT Applicant's Guide.

Docketed

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer J. Zahra	For/10/9/2000 by
Facsimile No. (41-22) 740.14.35	Telephone No. (41-22) 338.83.38	

Encl in pocket

CM

32313
PCT

PATENT COOPERATION TREATY

PCT

INFORMATION CONCERNING ELECTED OFFICES NOTIFIED OF THEIR ELECTION

(PCT Rule 61.3)

From the INTERNATIONAL BUREAU, BAKER BOTTS L.L.P.

To:

01 JAN 23 PM 12: 47

TANG, Henry
 Baker Botts, LLP
 30 Rockefeller Plaza
 New York, NY 10112-0228
 ETATS-UNIS D'AMERIQUE

Date of mailing (day/month/year)
 10 January 2001 (10.01.01)

Applicant's or agent's file reference
 32313-PCT

IMPORTANT INFORMATION

International application No. PCT/US00/04118	International filing date (day/month/year) 18 February 2000 (18.02.00)	Priority date (day/month/year) 19 February 1999 (19.02.99)
---	---	---

Applicant

THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK et al

1. The applicant is hereby informed that the International Bureau has, according to Article 31(7), notified each of the following Offices of its election:

AP : GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW

EP : AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE

National : AU, BG, CA, CN, CZ, DE, IL, JP, KP, KR, MN, NO, NZ, PL, RO, RU, SE, SK, US

2. The following Offices have waived the requirement for the notification of their election; the notification will be sent to them by the International Bureau only upon their request:

EA : AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

OA : BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

National : AE, AL, AM, AT, AZ, BA, BB, BR, BY, CH, CR, CU, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IN, IS, KE, KG, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MW, MX, PT, SD, SG, SI, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW

3. The applicant is reminded that he must enter the "national phase" before the expiration of 30 months from the priority date before each of the Offices listed above. This must be done by paying the national fee(s) and furnishing, if prescribed, a translation of the international application (Article 39(1)(a)), as well as, where applicable, by furnishing a translation of any annexes of the international preliminary examination report (Article 36(3)(b) and Rule 74.1).

Some offices have fixed time limits expiring later than the above-mentioned time limit. For detailed information about the applicable time limits and the acts to be performed upon entry into the national phase before a particular Office, see Volume II of the PCT Applicant's Guide.

The entry into the European regional phase is postponed until 31 months from the priority date for all States designated for the purposes of obtaining a European patent.

The International Bureau of WIPO
 34, chemin des Colombettes
 1211 Geneva 20, Switzerland

Faxsimile No. (41-22) 740.14.35

Authorized officer:

R. E. Stoffel

Telephone No. (41-22) 338.83.38

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To: HENRY TANG
BAKER BOTTS, LLP
30 ROCKEFELLER PLAZA
NEW YORK NY 10112-0228

D JUN 26 AM 11:45

PCT

TO

NOTIFICATION OF TRANSMITTAL OF
INTERNATIONAL PRELIMINARY
EXAMINATION REPORT

(PCT Rule 71.1)

Date of Mailing
(day/month/year)

22 JUN 2001

Applicant's or agent's file reference
32313-PCT

IMPORTANT NOTIFICATION

International application No.

International filing date (day/month/year)

Priority Date (day/month/year)

PCT/US00/04118

18 FEBRUARY 2000

19 FEBRUARY 1999

Applicant

THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.
4. REMINDER

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

PATRICK N. EDGWARD

Telephone No. (703) 308-6728

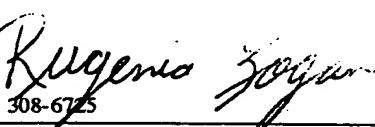
ATENT COOPERATION T
PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 32313-PCT	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US00/04118	International filing date (day/month/year) 18 FEBRUARY 2000	Priority date (day/month/year) 19 FEBRUARY 1999
International Patent Classification (IPC) or national classification and IPC IPC(7): GO6F 17/10, 17/27 AND 15/00 and US Cl.: 704/1, 9, 10; 707/530, 531, 532		
Applicant THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK		

<ol style="list-style-type: none"> This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36. This REPORT consists of a total of <u>3</u> sheets. <p><input type="checkbox"/> This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).</p> <p>These annexes consist of a total of <u>0</u> sheets.</p>
<ol style="list-style-type: none"> This report contains indications relating to the following items: <ul style="list-style-type: none"> I <input checked="" type="checkbox"/> Basis of the report II <input type="checkbox"/> Priority III <input type="checkbox"/> Non-establishment of report with regard to novelty, inventive step or industrial applicability IV <input type="checkbox"/> Lack of unity of invention V <input checked="" type="checkbox"/> Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement VI <input type="checkbox"/> Certain documents cited VII <input type="checkbox"/> Certain defects in the international application VIII <input type="checkbox"/> Certain observations on the international application

Date of submission of the demand 05 SEPTEMBER 2000	Date of completion of this report 11 FEBRUARY 2001
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Patrick N. Edouard  Telephone No. (703) 308-6725

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US00/04118

I. Basis of the report**1. With regard to the elements of the international application:*** the international application as originally filed the description:

pages 1-9

pages NONE

pages NONE

, filed with the demand

 the claims:

pages 10-14

, as originally filed

pages NONE

pages NONE

, filed with the demand

pages NONE

, filed with the letter of

 the drawings:

pages 1-5

, as originally filed

pages NONE

, filed with the demand

pages NONE

, filed with the letter of

 the sequence listing part of the description:

pages NONE

, as originally filed

pages NONE

, filed with the demand

pages NONE

, filed with the letter of

2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language _____ which is:

 the language of a translation furnished for the purposes of international search (under Rule 23.1(b)). the language of publication of the international application (under Rule 48.3(b)). the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).**3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:** contained in the international application in printed form. filed together with the international application in computer readable form. furnished subsequently to this Authority in written form. furnished subsequently to this Authority in computer readable form. The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished. The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.**4. The amendments have resulted in the cancellation of:** the description, pages NONE the claims, Nos. NONE the drawings, sheets/fig NONE**5. This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).******* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).******Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.**

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US00/04118

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**1. statement**

Novelty (N)	Claims 1-22	YES
	Claims <u>NONE</u>	NO
Inventive Step (IS)	Claims 1-22	YES
	Claims <u>NONE</u>	NO
Industrial Applicability (IA)	Claims 1-22	YES
	Claims <u>NONE</u>	NO

2. citations and explanations (Rule 70.7)

Claims 1-22 meet the criteria set out in PCT Article 33(2)-(4), because the prior art does not teach or fairly suggest a method/system/computer readable media for generating a summary of a plurality of related documents in a collection comprising performing phrase intersection analysis on the extracted phrases to generate a phrase intersection table; performing temporal porcessing on the phrases in the phrase intersection table and performing sentence generation using the phrase in the phrase intersection.

----- NEW CITATIONS -----
NONE

The demand must be filed directly with a competent International Preliminary Examining Authority or, if two or more Authorities are with the one chosen by the applicant, the full name or two-letter code of that Authority may be indicated by the applicant on the line

IPEA/ US

PCT

09/913745

CHAPTER II

DEMAND

under Article 31 of the Patent Cooperation Treaty:

The undersigned requests that the international application specified below be the subject of international preliminary examination according to the Patent Cooperation Treaty and hereby elects all eligible States (except where otherwise indicated).

For International Preliminary Examining Authority use only

Identification of IPEA

Date of receipt of DEMAND

Box No. I IDENTIFICATION OF THE INTERNATIONAL APPLICATION		Applicant's or agent's file reference 32312-PCT
International application No. PCT/US00/04118	International filing date (day/month/year) 18 February 2000 (18.02.00)	(Earliest) Priority date (day/month/year) 19 February 1999 (19.02.99)
Title of invention MULTI-DOCUMENT SUMMARIZATION SYSTEM AND METHOD		
Box No. II APPLICANT(S)		
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.) THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK 116th Street and Broadway New York, NY 10027 US		Telephone No.: Facsimile No.: Teleprinter No.:
State (that is, country) of nationality: US		State (that is, country) of residence: US
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.) MCKEOWN, KATHLEEN R. 20 Prospect Road Wayne, NJ 07470 US		
State (that is, country) of nationality: US		State (that is, country) of residence: US
Name and address: (Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.) BARZILAY, REGINA 548 Riverside Drive, Apt. 4B New York, NY 10027 US		
State (that is, country) of nationality: US		State (that is, country) of residence: US
<input type="checkbox"/> Further applicants are indicated on a continuation sheet.		

Box No. III AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE

The following person is agent common representative

and has been appointed earlier and represents the applicant(s) also for international preliminary examination.

is hereby appointed and any earlier appointment of (an) agent(s) /common representative is hereby revoked.

is hereby appointed, specifically for the procedure before the International Preliminary Examining Authority, in addition to the agent(s)/common representative appointed earlier.

Name and address: (*Family name followed by given name; for a legal entity, full official
The address must include postal code and name of country.*)

TANG, HENRY and
ACKERMAN, PAUL D.
Baker Botts LLP
30 Rockefeller Plaza
New York, NY 10112-0228
US

Telephone No.:
(212) 705-5000

Facsimile No.:
(212) 705-5020

Teleprinter No.:

Address for correspondence: Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

Box No. IV BASIS FOR INTERNATIONAL PRELIMINARY EXAMINATION**Statement concerning amendments:***

1. The applicant wishes the international preliminary examination to start on the basis of:

the international application as originally filed.

the description as originally filed
 as amended under Article 34

the claims as originally filed
 as amended under Article 19 (together with any accompanying statement)
 as amended under Article 34

the drawings as originally filed
 as amended under Article 34

2. The applicant wishes any amendment to the claims under Article 19 to be considered as reversed.

3. The applicant wishes the start of the international preliminary examination to be postponed until the expiration of 20 months from the priority date unless the International Preliminary Examining Authority receives a copy of any amendments made under Article 19 or a notice from the applicant that he does not wish to make such amendments (Rule 69.1(d)). (*This check-box may be marked only where the time limit under Article 19 has not yet expired.*)

* Where no check-box is marked, international preliminary examination will start on the basis of the international application as originally filed or, where a copy of amendments to the claims under Article 19 and/or amendments of the international application under Article 34 are received by the International Preliminary Examining Authority before it has begun to draw up a written opinion or the international preliminary examination report, as so amended.

Language for the purposes of international preliminary examination: English

which is the language in which the international application was filed.
 which is the language of a translation furnished for the purposes of international search.
 which is the language of publication of the international application.
 which is the language of the translation (to be) furnished for the purposes of international preliminary examination.

Box No. V ELECTION OF STATES

The applicant hereby elects all eligible States (*that is, all States which have been designated and which are bound by Chapter II of the PCT*)

excluding the following States which the applicant wishes not to elect:

Box No. VI CHECK LIST

The demand is accompanied by the following elements, in the language referred to in Box No. IV, for the purposes of international preliminary examination:

			For International Preliminary Examining Authority use only
			received not received
1.	translation of international application	: sheets	<input type="checkbox"/> <input checked="" type="checkbox"/>
2.	amendments under Article 34	: sheets	<input type="checkbox"/> <input checked="" type="checkbox"/>
3.	copy (or where required, translation) of amendments under Article 19	: sheets	<input type="checkbox"/> <input checked="" type="checkbox"/>
4.	copy (or, where required, translation) of statement under Article 19	: sheets	<input type="checkbox"/> <input checked="" type="checkbox"/>
5.	letter	: sheets	<input type="checkbox"/> <input checked="" type="checkbox"/>
6.	other (specify)	: sheets	<input type="checkbox"/> <input checked="" type="checkbox"/>

The demand is also accompanied by the item(s) marked below:

1. <input checked="" type="checkbox"/> fee calculation sheet	4. <input type="checkbox"/> statement explaining lack of signature
2. <input type="checkbox"/> separate signed power of attorney	5. <input type="checkbox"/> nucleotide and or amino acid sequence listing in computer readable form
3. <input type="checkbox"/> copy of general power of attorney; reference number, if any:	6. <input checked="" type="checkbox"/> other (specify): Transmittal Letter

Box No. VII SIGNATURE OF APPLICANT, AGENT OR COMMON REPRESENTATIVE

Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the demand).

Paul D. Ackerman (Agent)

For International Preliminary Examining Authority use only

1. Date of actual receipt of DEMAND:		
2. Adjusted date of receipt of demand due to CORRECTIONS under Rule 60.1(b):		
3. <input type="checkbox"/> The date of receipt of the demand is AFTER the expiration of 19 months from the priority date and item 4 or 5, below, does not apply.	<input type="checkbox"/> The applicant has been informed accordingly.	
4. <input type="checkbox"/> The date of receipt of the demand is WITHIN the period of 19 months from the priority date as extended by virtue of Rule 80.5.		
5. <input type="checkbox"/> Although the date of receipt of the demand is after the expiration of 19 months from the priority date, the delay in arrival is EXCUSED pursuant to Rule 82.		

For International Bureau use only

Demand received from IPEA on:

PCT

FEE CALCULATION SHEET

Annex to the Demand for international preliminary examination

International application No.	PCT/US00/04118	For International Preliminary Examining Authority use only
Applicant's or agent's file reference	32313-PCT	Date stamp of the IPEA
<p>Applicant THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK</p>		
Calculation of prescribed fees		
1. Preliminary examination fee	490.00	P
2. Handling fee (<i>Applicants from certain States are entitled to a reduction of 75% of the handling fee. Where the applicant is (or all applicants are) so entitled, the amount to be entered at H is 25% of the handling fee.</i>)	153.00	H
3. Total of prescribed fees Add the amounts entered at P and H and enter total in the TOTAL box	643.00	TOTAL
Mode of Payment		
<input type="checkbox"/> authorization to charge deposit account with the IPEA (see below)	<input type="checkbox"/> cash	
<input checked="" type="checkbox"/> cheque	<input type="checkbox"/> revenue stamps	
<input type="checkbox"/> postal money order	<input type="checkbox"/> coupons	
<input type="checkbox"/> bank draft	<input type="checkbox"/> other (<i>specify</i>): _____	

Deposit Account Authorization (<i>this mode of payment may not be available at all IPEAs</i>)		
The IPEA/ <u>US</u>	<input type="checkbox"/>	is hereby authorized to charge the total fees indicated above to my deposit account.
<input checked="" type="checkbox"/> (<i>this check-box may be marked only if the conditions for deposit accounts of the IPEA so permit</i>) is hereby authorized to charge any deficiency or credit any overpayment in the total fees indicated above to my deposit account.		
02-4377	5 September 2000	
Deposit Account Number	Date (day/month/year)	Signature

TENT COOPERATION TREATY

09/913745

32313
PCT

From the INTERNATIONAL BUREAU

PCT**NOTIFICATION OF RECEIPT OF RECORD COPY**

(PCT Rule 24.2(a))

To:

BAKER BOTTS L.L.P.

00 JUN 12 PM 3:06

TANG, Henry
 Baker Botts, LLP
 30 Rockefeller Plaza
 New York, NY 10112-0228
 ETATS-UNIS D'AMERIQUE

[Handwritten signatures and initials over the address]

Date of mailing (day/month/year) 26 May 2000 (26.05.00)	IMPORTANT NOTIFICATION
--	-------------------------------

Applicant's or agent's file reference 32313-PCT	International application No. PCT/US00/04118
--	---

The applicant is hereby notified that the International Bureau has received the record copy of the international application as detailed below.

Name(s) of the applicant(s) and State(s) for which they are applicants:

THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK (for all designated States except US)
 MCKEOWN, Kathleen, R. et al (for US)

International filing date : 18 February 2000 (18.02.00)
 Priority date(s) claimed : 19 February 1999 (19.02.99)
 Date of receipt of the record copy by the International Bureau : 08 May 2000 (08.05.00)

List of designated Offices :

AP : GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW
 EA : AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
 EP : AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE
 OA : BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
 National : AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland	Authorized officer:
---	---------------------

Facsimile No. (41-22) 740.14.35	Telephone No. (41-22) 338.83.38
---------------------------------	---------------------------------


cm

Continuation of Form PCT/IB/301

NOTIFICATION OF RECEIPT OF RECORD COPY

Date of mailing (day/month/year) 26 May 2000 (26.05.00)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference 32313-PCT	International application No. PCT/US00/04118

ATTENTION

The applicant should carefully check the data appearing in this Notification. In case of any discrepancy between these data and the indications in the international application, the applicant should immediately inform the International Bureau.

In addition, the applicant's attention is drawn to the information contained in the Annex, relating to:

- time limits for entry into the national phase
- confirmation of precautionary designations
- requirements regarding priority documents

A copy of this Notification is being sent to the receiving Office and to the International Searching Authority.

INFORMATION ON TIME LIMITS FOR ENTERING THE NATIONAL PHASE

The applicant is reminded that the "national phase" must be entered before each of the designated Offices indicated in the Notification of Receipt of Record Copy (Form PCT/IB/301) by paying national fees and furnishing translations, as prescribed by the applicable national laws.

The time limit for performing these procedural acts is **20 MONTHS** from the priority date or, for those designated States which the applicant elects in a demand for international preliminary examination or in a later election, **30 MONTHS** from the priority date, provided that the election is made before the expiration of 19 months from the priority date. Some designated (or elected) Offices have fixed time limits which expire even later than 20 or 30 months from the priority date. In other Offices an extension of time or grace period, in some cases upon payment of an additional fee, is available.

In addition to these procedural acts, the applicant may also have to comply with other special requirements applicable in certain Offices. It is the applicant's responsibility to ensure that the necessary steps to enter the national phase are taken in a timely fashion. Most designated Offices do not issue reminders to applicants in connection with the entry into the national phase.

For detailed information about the procedural acts to be performed to enter the national phase before each designated Office, the applicable time limits and possible extensions of time or grace periods, and any other requirements, see the relevant Chapters of Volume II of the PCT Applicant's Guide. Information about the requirements for filing a demand for international preliminary examination is set out in Chapter IX of Volume I of the PCT Applicant's Guide.

GR and ES became bound by PCT Chapter II on 7 September 1996 and 6 September 1997, respectively, and may, therefore, be elected in a demand or a later election filed on or after 7 September 1996 and 6 September 1997, respectively, regardless of the filing date of the international application. (See second paragraph above.)

Note that only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination.

CONFIRMATION OF PRECAUTIONARY DESIGNATIONS

This notification lists only specific designations made under Rule 4.9(a) in the request. It is important to check that these designations are correct. Errors in designations can be corrected where precautionary designations have been made under Rule 4.9(b). The applicant is hereby reminded that any precautionary designations may be confirmed according to Rule 4.9(c) before the expiration of 15 months from the priority date. If it is not confirmed, it will automatically be regarded as withdrawn by the applicant. There will be no reminder and no invitation. Confirmation of a designation consists of the filing of a notice specifying the designated State concerned (with an indication of the kind of protection or treatment desired) and the payment of the designation and confirmation fees. Confirmation must reach the receiving Office within the 15-month time limit.

REQUIREMENTS REGARDING PRIORITY DOCUMENTS

For applicants who have not yet complied with the requirements regarding priority documents, the following is recalled.

Where the priority of an earlier national, regional or international application is claimed, the applicant must submit a copy of the said earlier application, certified by the authority with which it was filed ("the priority document") to the receiving Office (which will transmit it to the International Bureau) or directly to the International Bureau, before the expiration of 16 months from the priority date, provided that any such priority document may still be submitted to the International Bureau before that date of international publication of the international application, in which case that document will be considered to have been received by the International Bureau on the last day of the 16-month time limit (Rule 17.1(a)).

Where the priority document is issued by the receiving Office, the applicant may, instead of submitting the priority document, request the receiving Office to prepare and transmit the priority document to the International Bureau. Such request must be made before the expiration of the 16-month time limit and may be subjected by the receiving Office to the payment of a fee (Rule 17.1(b)).

If the priority document concerned is not submitted to the International Bureau or if the request to the receiving Office to prepare and transmit the priority document has not been made (and the corresponding fee, if any, paid) within the applicable time limit indicated under the preceding paragraphs, any designated State may disregard the priority claim, provided that no designated Office may disregard the priority claim concerned before giving the applicant an opportunity to furnish the priority document within a time limit which is reasonable under the circumstances.

Where several priorities are claimed, the priority date to be considered for the purposes of computing the 16-month time limit is the filing date of the earliest application whose priority is claimed.

09/03745 BAKER BOTTS L.L.P.

PATENT COOPERATION TREATY 00 AUG 15 PM 3:10

From the INTERNATIONAL SEARCHING AUTHORITY

TO *[Signature]*PCT *[Signature]*
*[Signature]*NOTIFICATION OF TRANSMITTAL OF
THE INTERNATIONAL SEARCH REPORT *[Signature]*
OR THE DECLARATION

(PCT Rule 44.1)

Date of Mailing
(day/month/year)**11 AUG 2000**

Applicant's or agent's file reference .32313-PCT	FOR FURTHER ACTION See paragraphs 1 and 4 below
International application No. PCT/US00/04118	International filing date (day/month/year) 18 FEBRUARY 2000
Applicant THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK	

1. The applicant is hereby notified that the international search report has been established and is transmitted herewith.

Filing of amendments and statement under Article 19:
The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

When? The time limit for filing such amendments is normally 2 months from the date of transmittal of the international search report; however, for more details, see the notes on the accompanying sheet.

Where? Directly to the International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland
Facsimile No.: (41-22) 740.14.35

Docketed

For *10/11/2000 By [Signature]*

For more detailed instructions, see the notes on the accompanying sheet.

2. The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

- the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.
- no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. Further action(s): The applicant is reminded of the following:

Shortly after 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in rules 90 bis 1 and 90 bis 3, respectively, before the completion of the technical preparations for international publication.

Within 19 months from the priority date, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later).

Within 20 months from the priority date, the applicant must perform the prescribed acts for entry into the national phase before all designated Offices which have not been elected in the demand or in a later election within 19 months from the priority date or could not be elected because they are not bound by Chapter II.

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer Patrick N. Edouard Telephone No. (703) 308-6725
Facsimile No. (703) 305-3230	<i>Joni Hill</i>

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 32313-PCT	FOR FURTHER ACTION	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
International application No. PCT/US00/04118	International filing date (day/month/year) 18 FEBRUARY 2000	(Earliest) Priority Date (day/month/year) 19 FEBRUARY 1999
Applicant THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 5 sheets.

It is also accompanied by a copy of each prior art document cited in this report.

1. Basis of the report

- a. With regard to the language, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.
 - the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).
- b. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of the sequence listing:
 - contained in the international application in written form.
 - filed together with the international application in computer readable form.
 - furnished subsequently to this Authority in written form.
 - furnished subsequently to this Authority in computer readable form.
 - the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
 - the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

2. Certain claims were found unsearchable (See Box I).

3. Unity of invention is lacking (See Box II).

4. With regard to the title,

- the text is approved as submitted by the applicant.
- the text has been established by this Authority to read as follows:

5. With regard to the abstract,

- the text is approved as submitted by the applicant.
- the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is Figure No. 1

- as suggested by the applicant.
- because the applicant failed to suggest a figure.
- because this figure better characterizes the invention.

None of the figures.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/04118

Box III TEXT OF THE ABSTRACT (Continuation of item 5 of the first sheet)

A summary for a collection of related documents can be generated by extracting phrases (100) from the documents which include common focus elements. phrase intersection analysis (120) is then performed on the extracted phrases (100) to generate a phrase intersection table, where identical or equivalent phrases are identified. temporal processing (140) on the phrases in the phrase intersection table is performed to remove ambiguous time reference and to sort the phrases in a temporal sequence. Sentence generation is then used to combine the phrases in the phrase intersection table into a coherent summary.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/04118

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : GO6F 17/10, 17/27, 15/00

US CL : 704/1, 9, 10; 707/530, 531, 532

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 704/1, 9, 10; 707/530, 531, 532

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	US 5,924,108 A (FEIN ET AL) 13 JULY 1999, ABSTRACT.	1-22
A	US 5,638,543 A (PEDERSEN ET AL) 10 JUNE 1997.	1-22
Y	US 5,077,668 A (DOI) 31 DECEMBER 1991, ABSTRACT.	1, 8 AND 16
A	US 5,848,191 A (CHEN ET AL) 08 DECEMBER 1998	1-22
A	US 4,965,763 A (ZAMORA) 23 OCTOBER 1990.	1-22
Y	US 5,838,323 A (ROSE ET AL) 17 NOVEMBER 1998, ABSTRACT.	1, 8 AND 16
A	US 5,297,027 A (MORIMOTO ET AL) 22 MARCH 1994.	1-22

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	
"B"	earlier document published on or after the international filing date	"X"
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"O"	document referring to an oral disclosure, use, exhibition or other means	"Y"
"P"	document published prior to the international filing date but later than the priority date claimed	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
	"A"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
03 JULY 2000	11 AUG 2000

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Patrick N. Edouard Telephone No. (703) 308-6725	Joni Hill
---	--	-----------

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/04118

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,689,716 A (CHEN ET AL) 18 NOVEMBER 1997, ABSTRACT.	1, 8 AND 16
A	US 5,778,397 A (KUPIEC ET AL) 07 JULY 1998.	1-22
A	US 5,384,703 A (WITHGOTT ET AL) 24 JANUARY 1995.	1-22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/04118

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST/EAST

search term: (summar\$ or abstract\$ or condens\$) same (document or text or information)

091913745

PATENT COOPERATION TREATY

PCT

REC'D 27 JUN 2001

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 32313-PCT	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/US00/04118	International filing date (day/month/year) 18 FEBRUARY 2000	Priority date (day/month/year) 19 FEBRUARY 1999
International Patent Classification (IPC) or national classification and IPC IPC(7): GO6F 17/10, 17/27 AND 15/00 and US Cl.: 704/1, 9, 10; 707/530, 531, 532		
Applicant THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

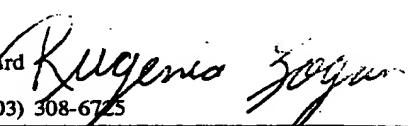
2. This REPORT consists of a total of 3 sheets.

This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 0 sheets.

3. This report contains indications relating to the following items:

- I Basis of the report
- II Priority
- III Non-establishment of report with regard to novelty, inventive step or industrial applicability
- IV Lack of unity of invention
- V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI Certain documents cited
- VII Certain defects in the international application
- VIII Certain observations on the international application

Date of submission of the demand 05 SEPTEMBER 2000	Date of completion of this report 11 FEBRUARY 2001
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer Patrick N. Edward Telephone No. (703) 308-6725 

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US00/04118

I. Basis of the report

1. With regard to the elements of the international application: *

 the international application as originally filed the description:

pages 1-9, as originally filed
 pages NONE, filed with the demand
 pages NONE, filed with the letter of _____

 the claims:

pages 10-14, as originally filed
 pages NONE, as amended (together with any statement) under Article 19
 pages NONE, filed with the demand
 pages NONE, filed with the letter of _____

 the drawings:

pages 1-5, as originally filed
 pages NONE, filed with the demand
 pages NONE, filed with the letter of _____

 the sequence listing part of the description:

pages NONE, as originally filed
 pages NONE, filed with the demand
 pages NONE, filed with the letter of _____

2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language _____ which is:

- the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- the language of publication of the international application (under Rule 48.3(b)).
- the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- contained in the international application in printed form.
- filed together with the international application in computer readable form.
- furnished subsequently to this Authority in written form.
- furnished subsequently to this Authority in computer readable form.
- The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

- the description, pages NONE
- the claims, Nos. NONE
- the drawings, sheets/fig NONE

5. This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).

**Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US00/04118

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**1. statement**

Novelty (N)	Claims <u>1-22</u>	YES
	Claims <u>NONE</u>	NO
Inventive Step (IS)	Claims <u>1-22</u>	YES
	Claims <u>NONE</u>	NO
Industrial Applicability (IA)	Claims <u>1-22</u>	YES
	Claims <u>NONE</u>	NO

2. citations and explanations (Rule 70.7)

Claims 1-22 meet the criteria set out in PCT Article 33(2)-(4), because the prior art does not teach or fairly suggest a method/system/computer readable media for generating a summary of a plurality of related documents in a collection comprising performing phrase intersection analysis on the extracted phrases to generate a phrase intersection table; performing temporal porcessing on the phrases in the phrase intersection table and performing sentence generation using the phrase in the phrase intersection.

----- NEW CITATIONS -----

NONE